Web Page Classification Using Convolutional Neural Network (CNN) Towards Eliminating Internet Addiction

Siti Hawa Apandi, Jamaludin Sallim, Rozlina Mohamed, Araby Madbouly
Faculty of Computing, College of Computing and Applied Sciences Universiti Malaysia Pahang, Pekan, Pahang, Malaysia

ABSTRACT
In the modern world, everyone has access to the internet as a source of information by surfing the web pages. The most popular web page surf is on Game and Online Video Streaming. Users who are spending too much time on these kinds of web pages may lead to a negative impact on Internet addiction. To overcome the internet addiction problem, access to Game and Online Video Streaming web pages needs to be restricted. Thus, a mechanism that can classify the category of the incoming web page based on the web page content is needed. This paper is proposing a web page classification model using a Convolutional Neural Network (CNN) to classify the web page, then identify whether it is a Game or Online Video Streaming based on the pattern of words in the word cloud image taken from the web page text content. The proposed web page classification model has achieved 82.22 % accuracy to detect the pre-classifled web pages.

KEYWORDS
Visualization; Scientific computing; Computational modeling; Web pages; Games; Streaming media;Predictive models

ACKNOWLEDGEMENTS

We would like to thank Dr. Azlee Zabidi, Senior Lecturer, Faculty of Computing, Universiti Malaysia Pahang for contributing very fruitful guidance to use MATLAB to develop web page classification model using Convolutional Neural Network (CNN). The work reported in this paper is supported by Universiti Malaysia Pahang Grant: RDU190310.